AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions of claims in the application:

Listing of Claims:

1. (Previously Presented) A system embodied on a computer-readable storage medium that facilitates decision tree learning, comprising:

a learning component that generates non-standardized data having a non-zero mean that relates to a split in a decision tree; and

a scoring component that assigns a score to the split as if the non-standardized data at a subset of leaves of the decision tree had been at least one of shifted or scaled, the non-standardized data is at least one of virtually shifted through omission of a matrix operation or virtually scaled through modification of a subset of elements relating to a covariance matrix, the score is at least one of stored on a computer-readable storage medium, displayed on a display device, employed by one or more processes executing on one or more processors, or transmitted between two or more processes executing on one or more processors.

2. (Previously Presented) The system of claim 1, further comprising a modification component that for a respective candidate split score, the data is modified by shifting or scaling the data and a new score is computed on the modified data.

3. (Original) The system of claim 1, further comprising an optimization component that analyzes the data and decides to treat the data as if it was: (1) shifted, (2) scaled, or (3) shifted and scaled.

4. (Original) The system of claim 1, the scoring component is employed for evaluating a data mining application.

5. (Original) The system of claim 1, the learning component processes continuous variable data or data subsets.

6. (Previously Presented) The system of claim 1, the scoring component generates evaluation data indicating how well a model predicts continuous target data and whether or not the model is a suitable predictor for the target data.

7. (Previously Presented) The system of claim 6, the evaluation data is employed by users or subsequent automated components when determining model performance or selecting between models or model subsets.

8. (Original) The system of claim 1, the scoring component includes at least one of a data sample processor, a scoring constant, a gamma function, a matrix value, a vector value, and a mean value for data or a data subset.

9. (Previously Presented) The system of claim 1, the scoring component computes a Bayesian linear regression score as:

$$score = \pi^{-n/2}\left(\frac{\nu}{\nu+n}\right)^{1/2}\frac{\Gamma\left(\frac{\alpha+n}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)}\left(\beta^{\frac{\alpha+r}{2}}\right)\frac{\left(\left|\mathbf{T}_{\mathbf{n}}^{\mathbf{TR}}\right|\right)^{-\left(\frac{\alpha+n}{2}\right)}}{\left(\left|\mathbf{T}_{\mathbf{n}}^{\mathbf{R}}\right|\right)^{-\left(\frac{\alpha-1+n}{2}\right)}},$$

$$\mathbf{T}_{\mathbf{n}} = \mathbf{T}_{\mathbf{0}} + \mathbf{S}_{\mathbf{n}} + \mathbf{U}_{\mathbf{n}}$$

$$\mathbf{U}_{\mathbf{n}} = \frac{\nu n}{\nu+n}(\overline{\mu}_0 - \overline{m}_n)(\overline{\mu}_0 - \overline{m}_n)'$$

$$\mathbf{S}_{\mathbf{n}} = \sum_{i=1}^{n}(\overline{x}_i - \overline{m}_n)(\overline{x}_i - \overline{m}_n)'$$

$$\overline{m}_n = \frac{1}{n}\sum_{i=1}^{n}\overline{x}_i$$

wherein μ represents a mean, α denotes a degree of freedom, β connotes a pre-defined constant, bold-face symbols denote square matrices, symbols with overlines denote (one dimensional) vectors, the ′ symbol denotes transpose, and $|\ |$ denotes determinant, n represents a number of records in the data, Γ is a gamma function satisfying Γ(x) = (x-1) Γ(x-1), $\overline{x}_i$ denotes a vector of values for relevant variables in an *ith* case in the data, the superscripts TR and R in $\mathbf{T}_{\mathbf{n}}^{\mathbf{TR}}$ and $\mathbf{T}_{\mathbf{n}}^{\mathbf{R}}$ denote that the matrices are defined with respect to target and regressor variables in a first case and regressor variables in a second case.

10. (Cancelled).

11. (Previously Presented) A system embodied on a computer-readable storage medium that facilitates data mining, comprising:

means for automatically generating a set of non-standardized data associated with a set or subset of data relating to a continuous variable, the non-standardized data associated with a split in a decision tree; and

means for automatically assigning a score to the split as if the non-standardized data were at least one of shifted or scaled, the non-standardized data is at least one of virtually shifted by omitting a matrix operation from automatically scoring the split or virtually scaled by modifying a subset of elements relating to a covariance matrix, the score is at least one of stored on a computer-readable storage medium, displayed on a display device, employed by one or more processes executing on one or more processors, or transmitted between two or more processes executing on one or more processors.

12. (Previously Presented) The system of claim 11, further comprising means for determining whether to perform the shifting operation and means for determining whether to perform the scaling operations.

13. (Previously Presented) The system of claim 11, further comprising means for shifting or scaling the set or subset of data relating to the continuous variable.

14. (Previously Presented) A computer-implemented method that facilitates decision tree learning, comprising:

determining whether to perform a virtual shifting operation on a non-standardized set of data with a non-zero mean associated with leaves of a decision tree;

determining whether to perform a virtual scaling operation on the non-standardized set of data; and

automatically assigning scores to the leaves based in part upon the determinations of whether to perform the virtual shifting and virtual scaling operations, the virtual shifting operation includes omitting a matrix operation from the assignment of scores and the virtual scaling operation includes modifying a subset of elements relating to a covariance matrix, the scores are at least one of stored on a computer-readable storage medium, displayed on a display device, employed by one or more processes executing on one or more processors, or transmitted between two or more processes executing on one or more processors.

15. (Previously Presented) The method of claim 14, further comprising performing at least one actual scaling or actual shifting operation on the non-standardized set of data.

16. (Original) The method of claim 14, further comprising processing a model in a form of a linear regression.

17. (Cancelled).

18. (Cancelled).

19. (Original) The method of claim 14, determining at least one constant value before assigning the scores.

20. (Original) The method of claim 19, the constant value relates to diagonal elements of a matrix and is assigned a value of about 0.01.

21. (Previously Presented) A computer readable storage medium that includes a tangible component that has a data structure stored thereon, comprising:

a first set of data fields describing a non-standardized set or subset of data relating to a continuous variable;

a second set of data fields describing a decision tree and associated branches; and

a third set of data fields describing a score for the branches, the score computed for the branches as if the non-standardized set or subset of data had been shifted or scaled, the non-standardized set or subset is at least one of virtually shifted by omission of a matrix operation from the computed score or virtually scaled by modification of a subset of elements relating to a covariance matrix, the score is at least one of stored on a computer-readable storage medium, displayed on a display device, employed by one or more processes executing on one or more processors, or transmitted between two or more processes executing on one or more processors.

22. (Original) The computer readable medium of claim 21, further comprising a data field to indicate at least one of a virtual shifting operation and a virtual scaling operation.

23. (Previously Presented) The computer readable medium of claim 21, further comprising a data field to indicate at least a portion of the non-standardized set or subset of data to be shifted or scaled.

24. (Cancelled)